

DOCUMENT RESUME

ED 452 565

CS 217 533

AUTHOR Greene, Beth G.
TITLE A Summary Report of the Instructional Effectiveness of the "Steck Vaughn Phonics Program": Level A--Units 3 and 4 & Level B--Units 2 and 3. Technical Report Number 105.
INSTITUTION Educational Research Inst. of America, Bloomington, IN.
PUB DATE 2001-06-01
NOTE 9p.
AVAILABLE FROM Educational Research Institute of America, PO Box 5875, Bloomington, IN 47408-5875.
PUB TYPE Reports - Research (143)
EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Instructional Effectiveness; *Phonics; Pilot Projects; Pretests Posttests; Primary Education; *Program Evaluation; *Reading
IDENTIFIERS California; Florida; Illinois; Subtests; Test Trials

ABSTRACT

This report describes a pilot study conducted to evaluate the instructional effectiveness of the Steck-Vaughn Phonics Program: Level A-Units 3 and 4 and Level B-Units 2 and 3. The Educational Research Institute of America (ERIA) identified volunteer teachers to participate, designed the assessments, and analyzed the collected data. Because the focus was on the program's primary levels, it was limited to Grades 1 and 2; Level A-Units 3 and 4 were used in first grade, while Level B-Units 2 and 3 were used in second grade. These units, selected by the teachers to best fit their schools' reading objectives at the time, included the teaching of eight initial consonants, six final consonants, short vowels, long vowels, and vowel diagraphs. A quasi-experimental pretest posttest design was used; 18 first and second grade teachers in California, Illinois, and Florida volunteered to participate. Approximately 400 students were included. Teachers administered the phonics tests developed by ERIA as pretests, taught the designated units from the Steck-Vaughn program, and administered the posttests after completing instruction. All assessments were scored by hand at ERIA. Results indicated that: (1) after about three weeks of instruction, two groups of first and second graders increased pretest scores to posttest scores a statistically significant amount on 5 out of 6 subtests and on total test scores for both the first and second grade assessments; (2) an analysis of students who scored lowest on the pretest indicated much greater improvement than was shown for the entire group--more than twice as great; (3) on the single subtest (second grade) where there was not a statistically significant improvement, there was evidence of a "ceiling" effect; and (4) all the teachers reported they were able to use the Steck-Vaughn Phonics Program without problems. Contains 5 tables of data. (NKA)

EDUCATIONAL RESEARCH INSTITUTE OF AMERICA

Post Office Box 5875 ♦ Bloomington ♦ Indiana ♦ 47408-5875

Technical Report Number 105
June 1, 2001

**A Summary Report of the Instructional Effectiveness
of the *Steck Vaughn Phonics Program*:
Level A—Units 3 and 4 & Level B—Units 2 and 3**

By Beth G. Greene

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

- ☐ This document has been reproduced as received from the person or organization originating it.
- ☐ Minor changes have been made to improve reproduction quality.
- Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

B. G. Greene

Directors:

Jennifer M. Conner
Ph. D. Indiana University

Beth G. Greene
Ph.D. New York University

Paul Lloyd
Ph.D. University of Adelaide,
South Australia

BEST COPY AVAILABLE

A Study of the Instructional Effectiveness of the Steck-Vaughn Phonics Program: Level A—Units 3 and 4 & Level B—Units 2 and 3

This report describes a pilot study that was conducted to evaluate the instructional effectiveness of the *Steck-Vaughn Phonics Program: Level A—Units 3 and 4 and Level B—Units 2 and 3*.

Background Information

The Steck-Vaughn Company engaged the *Educational Research Institute of America* (ERIA) to conduct an independent study to evaluate the effectiveness of typical units in its phonics program. The initial study selected two units in each of two levels: A and B (Grade 1 and Grade 2). ERIA identified volunteer teachers to participate in the study; designed the assessments used in the study; and analyzed the data that were collected. The study described in this report was conducted in the spring of the 2000-2001 school year.

Research Questions

The following research questions guided the design, procedures, and data analysis of the pilot study:

- *Is the Steck-Vaughn Phonics Program instructionally effective? Do selected units significantly increase students' recognition of initial and final consonants, short and long vowels, and vowel digraphs?*
- *Is the Steck-Vaughn Phonics Program instructionally effective with those students who score at lower levels on a phonics assessment? Do selected units significantly increase these lower performing students' recognition of initial and final consonants, short and long vowels, and vowel digraphs?*

Design and Procedures of the Study

Instructional Units Included in the Study

Because this study was focused on the primary levels of the program, it was limited to grades 1 and 2. Level A—Units 3 and 4 were used in first grade. Level B—Units 2 and 3 were used in second grade. These units were selected after consultation with the teachers about which units would best fit the reading objectives of schools at the time when the study was to be implemented. These units included the teaching of eight initial consonants, six final consonants, short vowels, long vowels, and vowel digraphs.

Design of the Study

A quasi-experimental pre-test-post-test design was used for the study. Eighteen first and second grade teachers in California, Illinois, and Florida volunteered to participate in the study. These teachers represent three different school districts. There were a total of approximately 400 students included in the study.

These volunteer teachers had not previously used the *Steck-Vaughn Phonics Program*. They administered the phonics tests developed by ERIA as pre-tests, taught the units designated from the *Steck-Vaughn Phonics Program*, and administered the post-tests after completing the instruction.

The Phonics Assessments

Experienced test developers at ERIA developed all of the assessments. The assessments for Level A (Grade 1) included *Beginning Consonants*, *Final Consonants*, and *Short Vowels* in three separated sub-tests. Each of the consonant sub-tests included 20 items; and the sub-test on vowels included 10 items.

The Level B (Grade 2) assessment covered *Short Vowels*, *Long Vowels*, and *Vowel Digraphs* in three sub-tests—each with 20 items.

The assessments sub-tests were developed to reflect the phoneme/grapheme relationships taught in the units used in the study. Also, the tests were designed to closely reflect the formats and emphases in several nationally administered standardized tests. All of the assessments were field tested for clarity for both teachers and students before they were used in the study.

The classroom teachers administered the assessments before and after the selected units were taught. ERIA prepared specific instructions for teachers as to how the tests were to be administered. Table 1 provides an outline of the assessments that were developed.

TABLE 1
Data-Collection Instruments

Assessment	Sub-tests
A	Matching initial consonant graphemes and phonemes: <i>b, j, k, n, p, s, t, w</i> (20 items)
A	Matching final consonant graphemes and phonemes: <i>b, k, p, n, s, t</i> (20 items)
A	Matching short vowel graphemes and phonemes: <i>o, i</i> (10 items)
A	Total score (50 items)
B	Matching short vowel phonemes and graphemes: <i>a, i, o, u</i> , and <i>e</i> (20 items)
B	Matching long vowel phonemes and graphemes: <i>a, i, o, u</i> , and <i>e</i> (20 items)
B	Matching vowel digraph phonemes and graphemes: <i>ai, ay, oa, oe, ow, ie, ea</i> , and <i>ee</i> (20 items)
B	Total score (60 items)

Instructional Procedures

Following the pre-tests, the teachers taught the selected units using the Steck-Vaughn Phonics Books. Teachers were instructed to follow the suggested teaching procedures outlined in the Teacher's Manual that accompanied each level of the program. ERIA established a toll-free telephone number for teachers to use if they had any questions regarding the instruction or the administration of the tests.

The time spent in completing the units varied from two to four weeks, depending on the schools in the three states and on how the instruction fit into the schedules in each classroom. Thus, it is assumed that the study reflects a range of schedules that are typically used when the program is adopted in schools.

Upon completion of the instruction using the selected units, students were administered the post-tests. All data collection-instruments were returned to ERIA where they were processed, scored, and analyzed.

All of the assessments were scored by hand at ERIA. Ten percent of the assessments were re-scored to assure the reliability of the scoring process. The reliability of the scoring process was .99. The data were analyzed by grade levels across school districts.

Two analyses were conducted. The first analyzed the pre- and post-test scores for all of the students who had taken both the pre- and post-tests. The second analysis included those students who scored the lowest on the pre-test. A paired sample t-test was used to statistically assess gain scores from pre-test to post-test for the all of the sub-tests as well as the total test scores. The results of the data analysis are presented within grades and by sub-tests in Tables 2 to 5 below.

Results

Table 2 shows the results of all the students in Grade 1 for whom pre- and post-test scores were available.

TABLE 2
Level A (Grade 1-All Students, N=173)

Sub-test	Pre-test Mean	Pre-test SD	Post-test Mean	Post-test SD	Diff.	t-test	Signif.
Beginning Consonants (20 items)	19.16	1.92	19.78	1.17	+.62	.73	<.001
Final Consonants (20 items)	18.32	2.65	19.62	1.64	+1.3	7.00	<.001
Short Vowels (10 items)	8.15	2.03	9.34	1.25	+1.19	8.24	<.001
Total (50 items)	45.65	5.28	48.74	1.8	+3.09	9.88	<.001

A second analysis was conducted by selecting all of those students whose pre-test total scores were 45 or less (out of a possible 50 points) These results are shown in Table 3.

TABLE 3
Level A (Grade-1 Low Scoring Students, N=48)

Sub-test	Pre-test Mean	Pre-test SD	Post-test Mean	Post-test SD	Diff.	t-test	Signif.
Beginning Consonants (20 items)	17.98	3.2	19.35	1.17	+1.37	2.94	<.001
Final Consonants (20 items)	15.98	4.0	19.0	2.8	+3.02	5.51	<.001
Short Vowels (10 items)	5.91	2.3	8.8	.40	+2.89	9.14	<.001
Total (50 items)	39.88	7.03	47.08	4.83	+7.02	9.42	<.001

Grade One Total Group Results: Of the approximately 200 students who were enrolled in the nine first grade classes, 173 had complete pre- and post-test scores. The mean scores for these 173 students for all of the sub-tests and the total test increased from pre-test to post-test. Moreover, all of the increases were statistically significant. These increases seem remarkable considering that the study was conducted in a relatively short period of time with *teachers who received no special training in the implementation of the program.*

It should be noted that the average student scores on the pre-test were quite high. On the Beginning Consonants sub-test, for example, the average pre-test score was 19.16 out of a possible 20 points. Despite this very high initial score, the increase on the post-test scores was statistically significant, indicating that the increase was consistent across all students.

Grade One Low Scoring Group Results: The analysis of students whose average scores on the pre-test were 45 or less provides an indication as to how much these lower achieving students improved. There was a total of 48 students who completed both the pre- and the post-tests and whose average scores were 45 or less. Their raw score increases were more than twice as large as those for the total group. All of the increases were statistically significant. The total raw score increase was over seven raw score points, an increase that is quite large for a fifty item test.

Table 4 shows the results of all the students in Grade 2 for whom pre- and post-test scores were available.

TABLE 4
Level B (Grade 2—All Students, N=161)

Sub-test	Pre-test Mean	Pre-test SD	Post-test Mean	Post-test SD	Diff.	t-test	Signif.
Short Vowels (20 items)	18.47	2.28	18.64	2.17	+.17	1.06	Non-Sig.
Long Vowels (20 items)	14.62	4.35	16.18	3.88	+1.56	5.10	<.001
Vowel Digraphs (20 items)	12.24	5.45	14.70	4.15	+2.46	8.31	<.001
Total (60 items)	45.42	10.09	49.44	8.51	+4.02	8.22	<.001

A second analysis was conducted by selecting all of those students whose pre-test total scores were 40 or less (out of a possible 60 points). These results are shown in Table 3.

TABLE 5
Level B (Grade 2—Low Scoring Students, N=52)

Sub-test	Pre-test Mean	Pre-test SD	Post-test Mean	Post-test SD	Diff.	t-test	Signif.
Short Vowels (20 items)	16.69	3.17	17.30	2.98	+.61	1.43	Non-Sig.
Long Vowels (20 items)	9.7	3.5	13.46	3.86	+3.76	5.84	<.001
Vowel Digraphs (20 items)	6.63	3.40	11.06	3.48	+4.43	10.18	<.001
Total (60 items)	33.06	5.28	41.69	6.90	+8.63	11.19	<.001

Grade Two Total Group Results: Of the approximately 200 students who were enrolled in the nine second grade classes, 161 had complete pre- and post-test scores. As with the first grade results, the mean scores increased for all of the sub-tests and the total test from pre-test to post-test. Moreover, all of the increases, except for the increase on the *Short Vowels* sub-test, were statistically significant. Again it should be pointed out that the teachers had no special training in the teaching of phonics and the study was conducted in relatively short period of time.

The ceiling effect (students scoring very high on the pre-test and thus not being able to show much improvement on the post-test) was especially evident for the *Short Vowels* sub-test where the pre-test average score was 18.47 out of a possible 20 points. The most the pre-test scores could improve (on average) was 1.53 points, and this would be possible only if every student had every answer correct on the post-test. Despite the ceiling effect on this sub-test, the scores did improve, although due to the high average score on the pre-test, the increase was slight.

On the sub-tests where there were not such high pre-test score averages, and thus greater room to demonstrate improvement, the post-test scores were significantly higher.

Grade Two Low Scoring Group Results: The analysis of students whose average scores on the pre-test were 40 or less provides an indication as to how much these lower achieving students improved. There was a total of 52 students who completed both the pre- and the post-tests and whose average scores were 40 or less. As with the first grade pre-test low scoring group, the second grade low scoring group increased their scores approximately twice as much as did the total group. All of the increases were statistically significant except for the *Short Vowels* where the increase was positive but not statistically significant. The total raw score increase was over eight raw score points, an increase that is quite large for a sixty item test.

Summary of Results

- After approximately three weeks of instructions, two groups of first and second grade students increased pre-test scores to post-test scores a statistically significant amount on 5 out of 6 sub-tests and on total test scores for both the first and second grade assessments. The increase at Grade 2 (using Level B) was greater than the increase at Grade 1 (using Level A).
- An analysis of students who scored lowest on the pre-test indicated a much greater improvement than was shown for the entire group. For both the low scoring first graders and the low scoring second graders, the improvement was more than twice as great. These results indicate two things. First, if students score at very high levels on a pre-test, there is little room for improvement. Second, low scoring students improved at a very significant rate on a skill area that many reading educators believe is important to overall reading development.

Steck-Vaughn Phonics Program Study

- On the single sub-test (second grade) where there was not a statistically significant improvement, there was evidence of a *ceiling* effect. Many students scored very high on this pre-test, leaving little room to register improvement.
- All of the teachers reported that they were able to use the Steck-Vaughn Phonics Program without problems.